

```

<!--StartFragment-->RESULT 1
SART3_HUMAN
ID      SART3_HUMAN              Reviewed;          963 AA.
AC      Q15020; Q2M2H0; Q58F06; Q8IUS1; Q96J95;
DT      07-FEB-2006, integrated into UniProtKB/Swiss-Prot.
DT      01-NOV-1996, sequence version 1.
DT      21-AUG-2007, entry version 57.
DE      Squamous cell carcinoma antigen recognized by T-cells 3 (SART-3)
DE      (hSART-3) (Tat-interacting protein of 110 kDa) (Tip110).
GN      Name=SART3; Synonyms=KIAA0156, TIP110;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
OC      Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND SUBCELLULAR LOCATION.
RX      PubMed=10463607;
RA      Yang D., Nakao M., Shichijo S., Sasatomi T., Takasu H., Matsumoto H.,
RA      Mori K., Hayashi A., Yamana H., Shirouzu K., Itoh K.;
RT      "Identification of a gene coding for a protein possessing shared tumor
RT      epitopes capable of inducing HLA-A24-restricted cytotoxic T
RT      lymphocytes in cancer patients.";
RL      Cancer Res. 59:4056-4063(1999).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2), TISSUE SPECIFICITY,
RP      SUBCELLULAR LOCATION, FUNCTION, AND INTERACTION WITH TAT.
RC      TISSUE=Fetal brain;
RX      MEDLINE=22075130; PubMed=11959860; DOI=10.1074/jbc.M200773200;
RA      Liu Y., Li J., Kim B.O., Pace B.S., He J.J.;
RT      "HIV-1 Tat protein-mediated transactivation of the HIV-1 long terminal
RT      repeat promoter is potentiated by a novel nuclear Tat-interacting
RT      protein of 110 kDa, Tip110.";
RL      J. Biol. Chem. 277:23854-23863(2002).
RN      [3]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC      TISSUE=Bone marrow;
RX      MEDLINE=96127530; PubMed=8590280; DOI=10.1093/dnares/2.4.167;
RA      Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;
RT      "Prediction of the coding sequences of unidentified human genes. IV.
RT      The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT      analysis of cDNA clones from human cell line KG-1.";
RL      DNA Res. 2:167-174(1995).
RN      [4]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 3).
RC      TISSUE=Brain, Eye, Skin, and Uterus;
RX      PubMed=15489334; DOI=10.1101/gr.2596504;
RG      The MGC Project Team;
RT      "The status, quality, and expansion of the NIH full-length cDNA
RT      project: the Mammalian Gene Collection (MGC).";
RL      Genome Res. 14:2121-2127(2004).
RN      [5]
RP      PHOSPHORYLATION [LARGE SCALE ANALYSIS] AT SER-10 AND SER-16, AND MASS
RP      SPECTROMETRY.
RC      TISSUE=Epithelium;
RX      PubMed=17081983; DOI=10.1016/j.cell.2006.09.026;
RA      Olsen J.V., Blagoev B., Gnäd F., Macek B., Kumar C., Mortensen P.,
RA      Mann M.;
RT      "Global, in vivo, and site-specific phosphorylation dynamics in
RT      signaling networks.";
RL      Cell 127:635-648(2006).

```

```

RN      [6]
RP      PHOSPHORYLATION [LARGE SCALE ANALYSIS] AT SER-852, AND MASS
RP      SPECTROMETRY.
RX      PubMed=17525332; DOI=10.1126/science.1140321;
RA      Matsuoka S., Ballif B.A., Smogorzewska A., McDonald E.R. III,
RA      Hurov K.E., Luo J., Bakalarski C.E., Zhao Z., Solimini N.,
RA      Lerenthal Y., Shiloh Y., Gygi S.P., Elledge S.J.;
RT      "ATM and ATR substrate analysis reveals extensive protein networks
RT      responsive to DNA damage.";
RL      Science 316:1160-1166(2007).
CC      -!- FUNCTION: Regulates Tat transactivation activity through direct
CC      interaction. May be a cellular factor for HIV-1 gene expression
CC      and viral replication.
CC      -!- SUBUNIT: Interacts with HIV-1 Tat.
CC      -!- SUBCELLULAR LOCATION: Cytoplasm. Nucleus speckle. Note=Localized
CC      in speckles. Expressed in the nucleus of all of the malignant
CC      tumor cell lines tested and the majority of cancer tissues with
CC      various histologies, including squamous cell carcinomas (SCC),
CC      adenocarcinomas, melanomas and leukemias cells. However, this
CC      protein is undetectable in the nucleus of any cell lines of
CC      nonmalignant cells or normal tissues, except for the testis.
CC      Expressed in the cytoplasm of all the proliferating cells,
CC      including normal and malignant cells, but not in normal tissues,
CC      except for the testis and the fetal liver.
CC      -!- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=3;
CC      Name=1;
CC      IsoId=Q15020-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=Q15020-2; Sequence=VSP_017250, VSP_017251;
CC      Note=No experimental confirmation available;
CC      Name=3;
CC      IsoId=Q15020-3; Sequence=VSP_017248, VSP_017249;
CC      Note=No experimental confirmation available;
CC      -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC      -!- PTM: Phosphorylated upon DNA damage, probably by ATM or ATR.
CC      -!- SIMILARITY: Contains 8 HAT repeats.
CC      -!- SIMILARITY: Contains 2 RRM (RNA recognition motif) domains.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; AF387506; AAK69347.1; -; mRNA.
DR      EMBL; AB020880; BAA78384.1; -; mRNA.
DR      EMBL; D63879; BAA09929.1; -; mRNA.
DR      EMBL; BC032601; AAH32601.1; -; mRNA.
DR      EMBL; BC041638; AAH41638.1; -; mRNA.
DR      EMBL; BC093784; AAH93784.1; -; mRNA.
DR      EMBL; BC103706; AAI03707.1; -; mRNA.
DR      EMBL; BC111983; AAI11984.1; -; mRNA.
DR      UniGene; Hs.584842; -.
DR      HSSP; Q14103; 1IQT.
DR      IntAct; Q15020; -.
DR      PeptideAtlas; Q15020; -.
DR      Ensembl; ENSG00000075856; Homo sapiens.
DR      HGNC; HGNC:16860; SART3.
DR      PharmGKB; PA34948; -.
DR      ArrayExpress; Q15020; -.
DR      GermOnline; ENSG00000075856; Homo sapiens.
DR      InterPro; IPR012677; a_b_plait_nuc_bd.
DR      InterPro; IPR003107; HAT.

```

DR InterPro; IPR008669; Lsm_interact.
 DR InterPro; IPR000504; RRM_RNP1.
 DR Gene3D; G3DSA:3.30.70.330; a_b_plait_nuc_bd; 2.
 DR Pfam; PF05391; Lsm_interact; 1.
 DR Pfam; PF00076; RRM_1; 2.
 DR SMART; SM00386; HAT; 7.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50102; RRM; 2.
 PE 1: Evidence at protein level;
 KW Alternative splicing; Coiled coil; Cytoplasm; Nucleus;
 KW Phosphorylation; Repeat; RNA-binding.
 FT CHAIN 1 963 Squamous cell carcinoma antigen
 FT recognized by T-cells 3.
 FT /FTid=PRO_0000223313.
 FT REPEAT 126 158 HAT 1.
 FT REPEAT 164 195 HAT 2.
 FT REPEAT 201 237 HAT 3.
 FT REPEAT 242 275 HAT 4.
 FT REPEAT 324 356 HAT 5.
 FT REPEAT 359 391 HAT 6.
 FT REPEAT 394 430 HAT 7.
 FT REPEAT 487 520 HAT 8.
 FT DOMAIN 704 782 RRM 1.
 FT DOMAIN 801 878 RRM 2.
 FT REGION 600 670 Required for nuclear localization.
 FT COILED 21 46 Potential.
 FT COILED 82 110 Potential.
 FT COILED 559 619 Potential.
 FT MOTIF 601 617 Nuclear localization signal (Potential).
 FT COMPBIAS 89 92 Poly-Glu.
 FT COMPBIAS 612 616 Poly-Lys.
 FT MOD_RES 10 10 Phosphoserine.
 FT MOD_RES 16 16 Phosphoserine.
 FT MOD_RES 852 852 Phosphoserine.
 FT VAR_SEQ 105 129 LSINVDYDYNCHVDLIRLLRLEGELT -> VGPGVGSGHLPV
 FT FQVLGSPCPGPPP (in isoform 3).
 FT /FTid=VSP_017248.
 FT VAR_SEQ 130 963 Missing (in isoform 3).
 FT /FTid=VSP_017249.
 FT VAR_SEQ 351 364 SQYLDRLKVKDLV -> RSTTESKGFGFICT (in
 FT isoform 2).
 FT /FTid=VSP_017250.
 FT VAR_SEQ 365 963 Missing (in isoform 2).
 FT /FTid=VSP_017251.
 SQ SEQUENCE 963 AA; 109935 MW; 06B26CEB8B19102A CRC64;

 Query Match 100.0%; Score 963; DB 1; Length 963;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 963; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATAAETSASEPEAESKAGPKADGEEDEVKAAARRKVL\$RAVAAATYKTMGPAWDQQEE 60
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MATAAETSASEPEAESKAGPKADGEEDEVKAAARRKVL\$RAVAAATYKTMGPAWDQQEE 60

 Qy 61 GVSESDGDEYAMASSAESSPGEYEWYDEEEEEKNQLEIERLEEQL\$INVDYDYNCHVDLIR 120
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 61 GVSESDGDEYAMASSAESSPGEYEWYDEEEEEKNQLEIERLEEQL\$INVDYDYNCHVDLIR 120

 Qy 121 LLRLEGELTKVRMARQKMSEIFPLTEELWLEWLHDEISMAQDGLDREHVYDLFEKAVKDY 180
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 121 LLRLEGELTKVRMARQKMSEIFPLTEELWLEWLHDEISMAQDGLDREHVYDLFEKAVKDY 180

```
Qy      181 ICPNIWLEYGQYSVGGIGQKGGLEKVRVSFERALSSVGLHMTKGLALWEAYREFESAIVE 240
      |||
Db      181 ICPNIWLEYGQYSVGGIGQKGGLEKVRVSFERALSSVGLHMTKGLALWEAYREFESAIVE 240

Qy      241 AARLEKVHSLFRRQLAIPLYDMEATFAEYEEWSEDPIPESVIQNYNKALQQLEKYKPYEE 300
      |||
Db      241 AARLEKVHSLFRRQLAIPLYDMEATFAEYEEWSEDPIPESVIQNYNKALQQLEKYKPYEE 300

Qy      301 ALLQAEAPRLAEYQAYIDFEMKIGDPARIQLIFERALVENCLVPDLWIRYSQYLDRLQKV 360
      |||
Db      301 ALLQAEAPRLAEYQAYIDFEMKIGDPARIQLIFERALVENCLVPDLWIRYSQYLDRLQKV 360

Qy      361 KDLVLSVHNRAIRNCPWTVALWSRYLLAMERHGVHDHQVISVTFEKALNAGFIQATDYVEI 420
      |||
Db      361 KDLVLSVHNRAIRNCPWTVALWSRYLLAMERHGVHDHQVISVTFEKALNAGFIQATDYVEI 420

Qy      421 WQAYLDYLRRRVDFKQDSSKELEELRAAFTRALEYLKQEVEERFNESGDPSCVIMQNWAR 480
      |||
Db      421 WQAYLDYLRRRVDFKQDSSKELEELRAAFTRALEYLKQEVEERFNESGDPSCVIMQNWAR 480

Qy      481 IEARLCNNMQKARELWDSIMTRGNAKYANMWLEYYNLERAHGDTQHCRKALHRAVQCTSD 540
      |||
Db      481 IEARLCNNMQKARELWDSIMTRGNAKYANMWLEYYNLERAHGDTQHCRKALHRAVQCTSD 540

Qy      541 YPEHVCEVLLTMERTEGSLEDWDIAVQKTETRLARVNEQRMKAAEKEAALVQQEEEEKAEQ 600
      |||
Db      541 YPEHVCEVLLTMERTEGSLEDWDIAVQKTETRLARVNEQRMKAAEKEAALVQQEEEEKAEQ 600

Qy      601 RKRRAEKKALKKKKKIRGPEKRGADEDDEKEWGDDEEEQPSKRRRVENSIPAAGETQNV 660
      |||
Db      601 RKRRAEKKALKKKKKIRGPEKRGADEDDEKEWGDDEEEQPSKRRRVENSIPAAGETQNV 660

Qy      661 EVAAGPAGKCAAVDVEPPSKQKEKAASLKRDMPKVLHDSSKDSITVFVSNLPYSMQEPDT 720
      |||
Db      661 EVAAGPAGKCAAVDVEPPSKQKEKAASLKRDMPKVLHDSSKDSITVFVSNLPYSMQEPDT 720

Qy      721 KLRPLFEACGEVVQIRPIFSNRGDFRGYCYVEFKEEKSALQALEMDRKSVEGRPMFVSPC 780
      |||
Db      721 KLRPLFEACGEVVQIRPIFSNRGDFRGYCYVEFKEEKSALQALEMDRKSVEGRPMFVSPC 780

Qy      781 VDKSKNPDFKVFYRSTSLKHKLFISGLPFSCTKEELEEEICKAHGTVKDLRLVTNRAGKP 840
      |||
Db      781 VDKSKNPDFKVFYRSTSLKHKLFISGLPFSCTKEELEEEICKAHGTVKDLRLVTNRAGKP 840

Qy      841 KGLAYVEYENESQASQAVMKMDGMTIKENIIKVAISNPPQRKVPEKPETRKAPGGPMLLP 900
      |||
Db      841 KGLAYVEYENESQASQAVMKMDGMTIKENIIKVAISNPPQRKVPEKPETRKAPGGPMLLP 900

Qy      901 QTYGARGKGRTQLSLLPRALQRPSAAAPQAENGPAAAPAVAAPAATEAPKMSNADFALF 960
      |||
Db      901 QTYGARGKGRTQLSLLPRALQRPSAAAPQAENGPAAAPAVAAPAATEAPKMSNADFALF 960

Qy      961 LRK 963
      |||
Db      961 LRK 963
<!--EndFragment-->
```